Hindawi Publishing Corporation
EURASIP Journal on Wireless Communications and Networking
Volume 2009, Article ID 142374, 12 pages
doi:10.1155/2009/142374

Research Article

Compound Wiretap Channels

Yingbin Liang,[1] Gerhard Kramer,[2] H. Vincent Poor,[3] and Shlomo Shamai (Shitz)[4]

[1] *Department of Electrical Engineering, University of Hawaii, Honolulu, HI 96822, USA*
[2] *Department of Electrical Engineering, University of Southern California, Los Angeles, CA 90089, USA*
[3] *Department of Electrical Engineering, Princeton University, Princeton, NJ 08544, USA*
[4] *Department of Electrical Engineering, Technion-Israel Institute of Technology, Technion City, Haifa 32000, Israel*

Correspondence should be addressed to Yingbin Liang, yingbinl@hawaii.edu

Received 1 December 2008; Revised 5 August 2009; Accepted 24 August 2009

Recommended by Hesham El-Gamal

This paper considers the compound wiretap channel, which generalizes Wyner's wiretap model to allow the channels to the (legitimate) receiver and to the eavesdropper to take a number of possible states. No matter which states occur, the transmitter guarantees that the receiver decodes its message and that the eavesdropper is kept in full ignorance about the message. The compound wiretap channel can also be viewed as a multicast channel with multiple eavesdroppers, in which the transmitter sends information to all receivers and keeps the information secret from all eavesdroppers. For the discrete memoryless channel, lower and upper bounds on the secrecy capacity are derived. The secrecy capacity is established for the degraded channel and the semideterministic channel with one receiver. The parallel Gaussian channel is further studied. The secrecy capacity and the secrecy degree of freedom (*s.d.o.f.*) are derived for the degraded case with one receiver. Schemes to achieve the *s.d.o.f.* for the case with two receivers and two eavesdroppers are constructed to demonstrate the necessity of a prefix channel in encoder design. Finally, the multi-antenna (i.e., MIMO) compound wiretap channel is studied. The secrecy capacity is established for the degraded case and an achievable *s.d.o.f.* is given for the general case.

1. Introduction

The compound channel models transmission over a channel that may take a number of states and reliable communication needs to be guaranteed regardless of which state occurs. For example, this type of channel might arise in real-time wireless communications when the transmitter has no knowledge of the channel state, but zero performance outage needs to be guaranteed subject to a stringent delay constraint. In this paper, we are interested in the compound channel with an eavesdropper that receives outputs via a compound channel that may also take a number of states. Now the transmitter not only needs to guarantee reliable communication to the legitimate receiver, but also needs to prevent the information from being known by the eavesdropper. This is a generalization of Wyner's wiretap channel [1] to the case of multiple channel states.

We consider the situation in which the channel remains in the same state during the entire transmission, and the channel state is known at the corresponding receivers, but not at the transmitter. However, we note that having the channel state information at the receivers comes at no cost to the communication rate, because the channel states can be learned by the receivers at the beginning of transmission via training symbols whose length is negligible compared to the codeword length.

We can also interpret the compound wiretap channel as the *multicast channel with multiple eavesdroppers* (see Figure 1). In this case, the number of states to the receiver now becomes the number of receivers with each state corresponding to one receiver, and the number of states to the eavesdropper becomes the number of eavesdroppers with each state corresponding to one eavesdropper. The transmitter wishes to transmit information to all receivers and keep the information secret from all eavesdroppers. In this paper, we adopt this interpretation. From this viewpoint, the compound wiretap channel provides a general framework that includes a number of models studied previously as special cases. These models include the parallel wiretap channel with two eavesdroppers studied in [2, 3], the fading

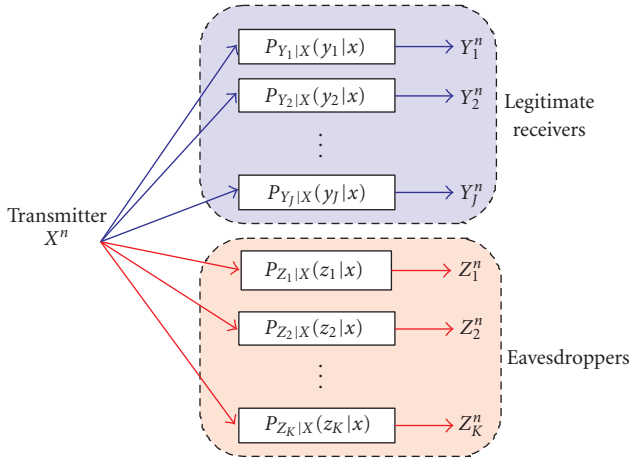

Figure 1: Compound wiretap channel.

wiretap channels with multiple eavesdroppers studied in [4], and the wiretap channel with multiple receivers studied in [5].

In this paper, we first study the discrete memoryless compound wiretap channel, for which we provide lower and upper bounds on the secrecy capacity. The lower bound indicates that the channel input scheme needs to balance the rates for all receiver-eavesdropper pairs, and hence none of them may achieve their best secrecy rate. We further establish the secrecy capacity for the degraded channel and the semideterministic channel with one receiver and multiple eavesdroppers.

We further study the parallel Gaussian compound wiretap channel, in which the channels to each receiver and to each eavesdropper are parallel Gaussian channels with multiple Gaussian subchannels. Channels of this type arise, for example, in wideband wireless communication systems such as frequency division multiplexing (FDM) systems in which transmission takes place over a number of frequency bands, and the eavesdroppers can tune their receivers to access some of these frequency bands. Understanding this channel is also important for studying the compound time-varying fading wiretap channels, as the parallel channel serves as a general model for the fading channel.

We first consider the degraded parallel Gaussian compound channel with one receiver and multiple eavesdroppers, for which we obtain the secrecy capacity. To further illustrate our results, we study the secrecy degree of freedom (*s.d.o.f.*), which characterizes how the secrecy capacity scales with log SNR. We show that the *s.d.o.f.* depends only on the total number of subchannels that the receiver accesses and the maximal number of subchannels that one eavesdropper can access. It is somewhat interesting that the *s.d.o.f.* does not depend on the total number of subchannels that all eavesdroppers can access and does not depend on the number of eavesdroppers either. We observe that there is a connection between the *s.d.o.f.* and secure network coding studied in [6]. However, the *s.d.o.f.* is defined for noisy Gaussian channels while secure network coding addresses deterministic networks.

We then study an example parallel Gaussian compound wiretap channel with two receivers and two eavesdroppers. For this channel, we propose three schemes. Scheme 1 is to map source information directly to Gaussian channel inputs, and this scheme is shown to be strictly suboptimal. Scheme 2 is to introduce a key random variable to randomize the source information, and this scheme achieves the *s.d.o.f.* Scheme 3 is to randomize the encoder by introducing a random prefix channel, and this scheme is also shown to achieve the *s.d.o.f.* This example channel demonstrates that randomization of either source information or encoder is necessary to achieve the *s.d.o.f.* for the parallel Gaussian compound channels.

We finally study the multiinput multioutput (MIMO) compound wiretap channel. We first provide the secrecy capacity for the degraded MIMO compound wiretap channel. We then study the general MIMO compound wiretap channel, for which we propose an input scheme and derive an achievable *s.d.o.f.* (a lower bound on the *s.d.o.f.*) based on this scheme. Comparing with the MIMO channel without eavesdroppers, the achievable *s.d.o.f.* of the MIMO compound wiretap channel is reduced by the maximal dimension of the projection of wiretap channel matrices on the vector space spanned by the eigenvectors corresponding to nonzero eigenvalues of channel matrices to the receiver.

We further note that after our conference publication [7] appeared with the results presented here, another upper bound on the secrecy capacity of the compound wiretap channel was derived in [8]. The secrecy capacity result for the parallel Gaussian compound wiretap channel was also extended to the nondegraded parallel Gaussian compound wiretap channel with one receiver and multiple eavesdroppers in [8]. We also refer the reader to [9] for a review of recent studies on compound wiretap channels.

The rest of the paper is organized as follows. In Section 2, we introduce the model of the compound wiretap channel. In Section 3, we present our results on the discrete memoryless compound wiretap channel. In Sections 4 and 5, we provide the results on the secrecy capacity and the *s.d.o.f.* for two cases of the parallel Gaussian compound wiretap channel. In Section 6, we provide our results on the MIMO compound wiretap channel. In the last section, we give concluding remarks.

2. Channel Model

We consider the following compound wiretap channel model.

Definition 1. The compound wiretap channel consists of one finite channel input alphabet $\mathcal{X}$, J finite channel output alphabets $\mathcal{Y}_1,\ldots,\mathcal{Y}_J$, K finite channel output alphabets $\mathcal{Z}_1,\ldots,\mathcal{Z}_K$, and a set of the transition probability distributions for one channel use

$$P_{Y_jZ_k|X}\left(y_j,z_k \mid x\right) \quad \text{for } j=1,\ldots,J,\ k=1,\ldots,K, \tag{1}$$

where $x \in \mathcal{X}$ is the channel input from the transmitter, $y_j \in \mathcal{Y}_j$ is the channel output at receiver j, and $z_k \in \mathcal{Z}_k$ is the channel output at eavesdropper k. The channel is memoryless across channel uses.

As the correlation between Y_j and Z_k does not affect the secrecy capacity (similar to [10, Lemma 1]), without loss of optimality, we assume a transition probability of the form $P_{Y_j|X}P_{Z_k|X}$ as shown in Figure 1.

Definition 2. A $(2^{nR}, n)$ code for the compound wiretap channel consists of the following:

(i) a message set: $\mathcal{W} = \{1, 2, \ldots, 2^{nR}\}$ with the message W uniformly distributed over $\mathcal{W}$;

(ii) an encoder $f\colon \mathcal{W} \to \mathcal{X}^n$ mapping each message $w \in \mathcal{W}$ to a codeword $x^n \in \mathcal{X}^n$;

(iii) J decoders $g_j\colon \mathcal{Y}_j^n \to \mathcal{W}^{(j)}$ for $j = 1, \ldots, J$, each mapping received sequence y_j^n to a message $\hat{w}^{(j)} \in \mathcal{W}$ for $j = 1, \ldots, J$.

The average block error probability for receiver j for $j = 1, \ldots, J$ is defined as

$$P_{e,j} = \frac{1}{2^{nR}} \sum_{w=1}^{2^{nR}} Pr\left\{\hat{w}^{(j)} \neq w\right\}. \tag{2}$$

The secrecy level of the message W at eavesdropper k for $k = 1, \ldots, K$ is defined by the following equivocation rate:

$$\frac{1}{n} H(W \mid Z_k^n). \tag{3}$$

A rate-equivocation pair (R, R_e) is *achievable* if there exists a sequence of $(2^{nR}, n)$ codes with the average error probabilities

$$P_{e,j}^{(n)} \longrightarrow 0 \quad \text{for } j = 1, \ldots, J \tag{4}$$

as n goes to infinity and with the equivocation rate satisfying

$$R_e \leq \lim_{n \to \infty} \frac{1}{n} H(W \mid Z_k^n) \quad \text{for } k = 1, \ldots, K. \tag{5}$$

In this paper, we are interested in the case of perfect secrecy, that is, $R = R_e$. A secrecy rate R is *achievable* if the rate-equivocation pair (R, R) is achievable. The *secrecy capacity* is defined to be the maximal achievable secrecy rate.

3. Discrete Memoryless Compound Wiretap Channels

In the following, we provide lower and upper bounds on the secrecy capacity of the compound wiretap channel.

Theorem 1. *The following secrecy rate is achievable for the compound wiretap channel:*

$$\begin{aligned} R &= \max\left[\min_j I\left(U; Y_j\right) - \max_k I(U; Z_k)\right] \\ &= \max \min_{j,k}\left[I\left(U; Y_j\right) - I(U; Z_k)\right], \end{aligned} \tag{6}$$

where U is an auxiliary random variable, and the maximum is taken over all distributions P_{UX} that satisfy the Markov chain relationships:

$$U \longrightarrow X \longrightarrow \left(Y_j, Z_k\right) \quad \text{for } j = 1, \ldots, J,\ k = 1, \ldots, K. \tag{7}$$

Proof. See Appendix A. □

Theorem 1 can be interpreted as a worst case result that is, the worst receiver and the best eavesdropper dominate the secrecy rate.

Theorem 2. *An upper bound on the secrecy capacity of the compound wiretap channel is given by*

$$\overline{R} = \min_{j,k} \max_{P_{UX} P_{Y_j} P_{Z_k|X}} \left[I\left(U; Y_j\right) - I(U; Z_k)\right], \tag{8}$$

where U is an auxiliary random variable whose joint distribution with X, Y_j, and Z_k factors was shown in (8).

Proof. It can be seen that the quantity

$$\max_{P_{UX} P_{Y_j Z_k|X}} \left[I\left(U; Y_j\right) - I(U; Z_k)\right] \tag{9}$$

in (8) is the secrecy capacity of the wiretap channel with the transition probability distribution $P_{Y_j Z_k|X}$ [11, Corollary 2]. But the secrecy capacity of the compound wiretap channel is less than the secrecy capacity of any receiver-eavesdropper pair. □

We note that it may not be possible to achieve the upper bound given in Theorem 2 in general. This is because the input scheme needs to balance the rates that can be achieved for all receiver-eavesdropper pairs, and consequently, none of them can achieve its best rate. This can also be seen from the achievable rate in (6). The input distribution P_{UX} that maximizes the minimum of the secrecy rates of all receiver-eavesdropper pairs may not be optimal for any single pair.

We now give an example channel in which the lower bound given in Theorem 1 can be shown to be the secrecy capacity. We say that the compound wiretap channel is *degraded* if the transition probability satisfies the Markov chain relationships:

$$X \longrightarrow Y_j \longrightarrow Z_k \tag{10}$$

for all $j = 1, \ldots, J$ and $k = 1, \ldots, K$. For the degraded compound wiretap channel, we have the following capacity theorem.

Theorem 3. *The secrecy capacity of the degraded compound wiretap channel is given by*

$$\begin{aligned} C &= \max_{P_X}\left[\min_j I\left(X; Y_j\right) - \max_k I(X; Z_k)\right] \\ &= \max_{P_X} \min_{j,k}\left[I\left(X; Y_j\right) - I(X; Z_k)\right]. \end{aligned} \tag{11}$$

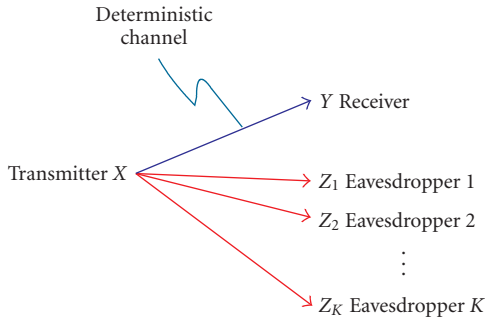

Figure 2: Semideterministic compound wiretap channel.

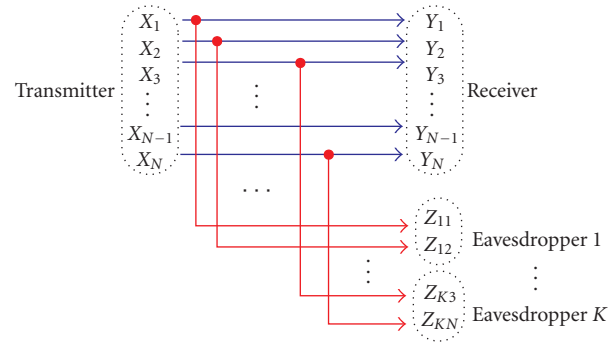

Figure 3: Parallel compound wiretap channel with one receiver and K eavesdroppers.

Proof. The achievability follows from Theorem 1 by setting $U = X$. The converse follows because for each (j,k) and an input distribution P_X, an upper bound

$$R_e \le I\left(X;Y_j\right) - I(X;Z_k) \tag{12}$$

can be derived as given in [1]. □

We next provide the secrecy capacity for the semideterministic compound wiretap channel, which has one receiver ($J = 1$) and K eavesdroppers. The channel from the transmitter to the receiver is a deterministic channel; that is, the transition probability distribution $P_{Y|X}$ takes on the values 0 or 1 only, where the output at the receiver is denoted by Y (see Figure 2).

Theorem 4. *The secrecy capacity of the semideterministic compound wiretap channel with $J = 1$ is given by*

$$C_s = \max_{P_X} \min_k H(Y \mid Z_k). \tag{13}$$

Proof. To prove the achievability, we apply (6) and obtain the following achievable rate:

$$R = \max \min_k [I(U;Y) - I(U;Z_k)], \tag{14}$$

where the maximum is taken over all distributions P_{UX} that satisfy the Markov chain relationship:

$$U \longrightarrow X \longrightarrow (Y,Z_k) \quad \text{for } k = 1,\ldots,K. \tag{15}$$

We further let $U = Y$. It is clear that this choice satisfies the previous Markov chain condition, and results in an achievable rate

$$R = \max \min_k H(Y \mid Z_k). \tag{16}$$

The converse is relegated to Appendix B. □

We note that the achievable scheme involves choosing an auxiliary random variable $U = Y$. This indicates that a prefix channel from U to the actual channel input X at the encoder is necessary to achieve the secrecy capacity.

4. Parallel Gaussian Compound Wiretap Channels: $J = 1$

In this section, we focus on the case in which $J = 1$ and $K > 1$, that is, one receiver and K eavesdroppers (see Figure 3). We further assume that the channel from the transmitter to the receiver is the parallel Gaussian channel with N independent subchannels, and the outputs of the subchannels at the receiver for one channel use are given by

$$Y_a = X_a + W_a \quad \text{for } a = 1,\ldots,N, \tag{17}$$

where $W_1,\ldots,W_a$ are independent Gaussian random variables with variances $w_1^2,\ldots,w_a^2$, and these noise variables are independent and identically distributed (i.i.d.) across channel uses. We note that for this model, $Y_1,\ldots,Y_N$ indicate the outputs at the receiver from the N subchannels, and do not indicate the outputs corresponding to different receivers. The channel input is subject to the average power constraint P, that is,

$$\frac{1}{n}\sum_{i=1}^{n}\sum_{a=1}^{N}[X_{ai}^2] \le P, \tag{18}$$

where i is the symbol time index. We assume that each eavesdropper can access some subchannels. On letting $\mathcal{A}_k \subseteq \{1,\ldots,N\}$ include all indices of the subchannels that eavesdropper k can access, the outputs at eavesdropper k are given by

$$Z_{ka} = X_a + V_{ka} \quad \text{for } a \in \mathcal{A}_k, \tag{19}$$

where V_{ka} for $a \in \mathcal{A}_k$ are independent Gaussian random variables with variances v_{ka}^2. We further assume that $v_{ka}^2 \ge w_a^2$ for all $a \in \mathcal{A}_k$, and hence the channel is *degraded*.

For the degraded parallel Gaussian compound wiretap channel, we have the following secrecy capacity.

Corollary 1. *The secrecy capacity of the degraded parallel Gaussian compound wiretap channel is given by*

$$C = \max_{\sum_{a=1}^{N} P_a \le P} \min_k \left[\sum_{a=1}^{N}\frac{1}{2}\log\left(1+\frac{P_a}{w_a^2}\right) - \sum_{a\in\mathcal{A}_k}\frac{1}{2}\log\left(1+\frac{P_a}{v_{ka}^2}\right)\right]. \tag{20}$$

Proof. The achievability follows from Theorem 3 by choosing $X = X_1, \ldots, X_N$ with independent components and each $X_a \in \mathcal{N}(0, P_a)$. The converse follows from [12, Theorem 2] by setting $R_0 = 0$ for each eavesdropper. □

We note that the parallel Gaussian compound wiretap channel is a more general model than the model in [3] in that the number of eavesdroppers is arbitrary, each eavesdropper may access an arbitrary number of subchannels, and the transmitter is allowed to allocate power among the subchannels to achieve better secrecy rate. We also note that the parallel Gaussian compound wiretap channel reduces to the Gaussian/fading wiretap channel with multiple eavesdroppers studied in [4] if there is only one subchannel.

We further note that after our conference publication [7] appeared with the results presented here, the secrecy capacity of the general (i.e., not necessarily degraded) parallel Gaussian compound wiretap channel with one receiver and multiple eavesdroppers has been obtained in [8]. We refer the reader to [8] for further details.

To gain further insight into the secrecy capacity, we consider the rate at which the secrecy capacity scales with logSNR. In particular, we define the secrecy degree of freedom (s.d.o.f.) as

$$s.d.o.f. = \lim_{\mathrm{SNR}\to\infty} \frac{C(\mathrm{SNR})}{(1/2)\log \mathrm{SNR}}, \tag{21}$$

where without loss of generality, we choose w_1^2 as the reference noise level and define $\mathrm{SNR} = P/(Nw_1^2)$. We refer to a lower bound on the *s.d.o.f.* as an *achievable s.d.o.f.*

Corollary 2. *Assume that the maximal number of subchannels that one eavesdropper can access is L. The secrecy degree of freedom of the degraded parallel Gaussian compound wiretap channel with one receiver is given by*

$$s.d.o.f. = N - L. \tag{22}$$

Proof. The achievability follows by applying Corollary 1 and choosing $P_a = P/N$ for $a = 1, \ldots, N$. The converse follows by considering only eavesdropper k that accesses L subchannels, that is, $|\mathcal{A}_k| = L$, and evaluating the first-order SNR expansion of the secrecy capacity. □

Remark 1. The *s.d.o.f.* depends only on the maximal number of subchannels that one eavesdropper can access and does not depend on the total number of subchannels that all eavesdroppers access. This is because the eavesdroppers do not cooperate with each other. This implies that, even if every subchannel is accessed by some eavesdropper, positive *s.d.o.f.* is still possible if none of the eavesdroppers accesses a full set of the subchannels. This can also be seen from the examples given in [7].

Remark 2. The *s.d.o.f.* does not depend on the number of eavesdroppers.

We note that the *s.d.o.f.* in Corollary 2 is similar to the secure rate given in [6, Theorem 2] for multicast networks based on network coding. However, we note that Corollary 2 is applicable for noisy Gaussian channels while the secure rate given in [6, Theorem 2] is derived for deterministic networks.

We also refer the reader to [7] for some example channels for which simple schemes were constructed to achieve the *s.d.o.f.*

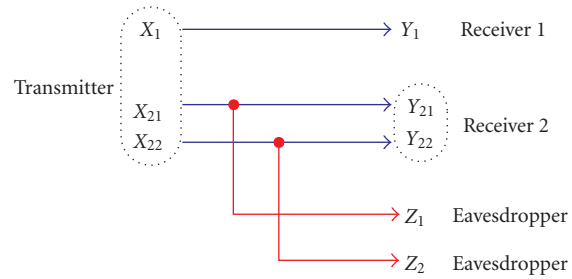

Figure 4: Parallel Gaussian compound wiretap channel example.

5. Parallel Gaussian Compound Wiretap Channels: $J > 1$

In this section, we study the parallel Gaussian compound wiretap channel, in which $J > 1$ and $K > 1$. We address optimal schemes that achieve the best secrecy rate scaling with SNR. For the sake of clarity of exposition on this issue, we study the simplest example when $J = 2$ and $K = 2$ to illustrate the key factors that affect optimal schemes.

Example 1. Consider the parallel Gaussian compound wiretap channel with $J = 2$ and $K = 2$ (see Figure 4). The channel output at receiver 1 is given by

$$Y_1 = X_1 + W_1, \tag{23}$$

where W_1 is a zero-mean Gaussian random variable with variance w_1^2. The channel outputs at receiver 2 are given by

$$Y_{21} = X_{21} + W_{21}, \quad Y_{22} = X_{22} + W_{22}, \tag{24}$$

where W_{21} and W_{22} are zero-mean independent Gaussian random variables with variances w_{21}^2 and w_{22}^2.

The outputs at the two eavesdroppers are given by

$$\begin{aligned} Z_1 &= X_{21} + V_1, \\ Z_2 &= X_{22} + V_2, \end{aligned} \tag{25}$$

where V_1 and V_2 are zero-mean independent Gaussian random variables with variances v_1^2 and v_2^2, respectively.

The channel input includes three components X_1, X_{21}, and X_{22}, and they are subject to an average power constraint P, that is,

$$\frac{1}{n}\sum_{i=1}^{n} \mathrm{E}[X_{1i}^2 + X_{21i}^2 + X_{22i}^2] \le P. \tag{26}$$

For this channel, we study the *s.d.o.f.*, for which we choose w_1^2 as the reference noise level and define $\mathrm{SNR} = P/w_1^2$.

An achievable rate follows from (6) and is given by

$$R = \max_{P_{UX}} \min\{I(U;Y_1) - I(U;Z_1), I(U;Y_1) - I(U;Z_2), I(U;Y_{21},Y_{22}) - I(U;Z_1), I(U;Y_{21},Y_{22}) - I(U;Z_2)\}. \tag{27}$$

In the following, we study three schemes, two of which are based on (27). It can be seen that a prefix channel $U \to X$ is necessary to achieve the optimal *s.d.o.f.*. For computational convenience, in the following we assume $w_1^2 = w_{21}^2 = w_{22}^2 = v_1^2 = v_2^2 = 1$. This assumption does not affect the *s.d.o.f.*, which we compute for each scheme.

Scheme 1. Choose $U = X = (X_1, X_{21}, X_{22})$ and $X_1 \sim \mathcal{N}(0, P_1), X_{21} \sim \mathcal{N}(0, P_{21})$, and $X_{22} \sim \mathcal{N}(0, P_{22})$ in (27). Based on these distributions, Scheme 1 achieves the following secrecy rate:

$$\begin{aligned} R &= \max_{P_1+P_{21}+P_{22}\le P} \min\{I(X_1;Y_1) - I(X_{21};Z_1), \\ &\qquad I(X_1;Y_1) - I(X_{22};Z_2), \\ &\qquad I(X_{21};Y_{21}) + I(X_{22};Y_{22}) - I(X_{21};Z_1), \\ &\qquad I(X_{21};Y_{21}) + I(X_{22};Y_{22}) - I(X_{22};Z_2)\} \\ &= \max_{P_1+P_{21}+P_{22}\le P} \min\left\{\frac{1}{2}\log(1+P_1) - \frac{1}{2}\log(1+P_{21}), \right. \\ &\qquad \frac{1}{2}\log(1+P_1) - \frac{1}{2}\log(1+P_{22}), \\ &\qquad \left. \frac{1}{2}\log(1+P_{22}), \frac{1}{2}\log(1+P_{21})\right\}. \end{aligned} \tag{28}$$

It can be seen that the optimal power allocation $(P_1^*, P_{21}^*, P_{22}^*)$ should result in four equal terms in the minimum in (28). Hence we obtain the following condition:

$$\frac{1}{2}\log(1+P_1^*) = \log(1+P_{21}^*) = \log(1+P_{22}^*). \tag{29}$$

Combining the preceding equation and the power constraint $P_1^* + P_{21}^* + P_{22}^* = P$, we obtain

$$\begin{aligned} P_1^* &= P - 2\sqrt{4+P} + 4, \\ P_{21}^* &= P_{22}^* = \sqrt{4+P} - 2. \end{aligned} \tag{30}$$

Substituting the optimal power allocation into (28), we obtain

$$R = \frac{1}{2}\log\left(\sqrt{4+P} - 1\right) \doteq \frac{1}{4}\log \text{SNR}, \tag{31}$$

where $(a \doteq b)$ denotes that $\lim_{P\to\infty}(a/b) = 1$.

Therefore, Scheme 1 achieves

$$s.d.o.f. = \frac{1}{2}. \tag{32}$$

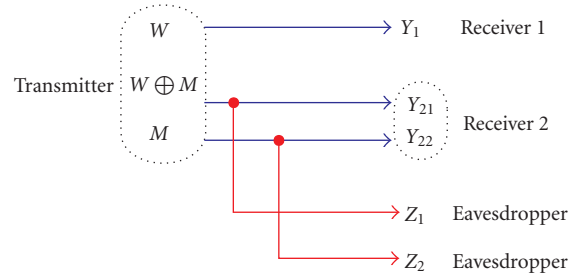

Figure 5: Illustration of Scheme 2.

Scheme 2. We choose a Gaussian input and allocate the source power equally for X_1, X_{21}, and X_{22}. Each subchannel can hence support the following rate:

$$R = \frac{1}{2}\log\left(1 + \frac{P}{3}\right). \tag{33}$$

Recall that the source message W is uniformly distributed over the set $\{0, \ldots, 2^{nR} - 1\}$. We generate a key random variable M that is independent of W and is also uniformly distributed over the set $\{0, \ldots, 2^{nR} - 1\}$. We transmit W over the channel $X_1 \to Y_1$ and transmit $W \oplus M$ and M over the channels $X_{21} \to Y_{21}$ and $X_{22} \to Y_{22}$, respectively (see Figure 5). It is clear that receiver 1 decodes W, and receiver 2 decodes $W \oplus M$ and M, and hence decodes W. For eavesdroppers 1 and 2, each obtains either $W \oplus M$ or M, both of which are independent of W. Hence eavesdroppers 1 and 2 do not get any information about W, and perfect secrecy is achieved. It is clear that this scheme achieves

$$s.d.o.f. = 1. \tag{34}$$

This is clearly the largest achievable *s.d.o.f.*, because the maximal degree of freedom achievable for receiver 1 is 1.

We note that Scheme 2 introduces randomness into the information source to achieve secrecy. Interestingly, Scheme 2 can be interpreted as turning the channel into a state dependent wiretap channel as studied in [13]. The key random variable M in Scheme 2 now corresponds to the channel state, which is known to the transmitter only. As shown in [13], the state variable helps improving the secrecy rate.

As remarked in Section 4, Scheme 2 for the noisy Gaussian channel is similar to the scheme designed for deterministic wiretap network models in [6]. More recently, deterministic network models have been proposed and studied (see, e.g., [14]) to obtain sufficiently accurate performance for Gaussian networks. It is hence interesting to apply this approach to study the secrecy capacity or *s.d.o.f.* for the Gaussian or other noisy wiretap networks. The key step is to come up with deterministic models that approximate the performance (e.g., in terms of *s.d.o.f.*) of noisy wiretap networks, and whose secrecy capacity can be determined easily.

Scheme 2 also suggests that Scheme 1 is strictly suboptimal. It is then natural to ask if we can modify Scheme 1 by

defining the auxiliary random variable U in (27) properly to achieve the optimal *s.d.o.f.* We hence propose the following Scheme 3.

Scheme 3. Choose $U = (X_1, X_{21} + X_{22})$ and $X_1 \sim \mathcal{N}(0, P/3), X_{21} \sim \mathcal{N}(0, P/3)$, and $X_{22} \sim \mathcal{N}(0, P/3)$ in (27). It is clear that the above choice of U satisfies the Markov chain relationship $U \to X \to (YZ)$ and is hence valid. The achievable secret rate under this scheme is given by

$$\begin{aligned} R = \min\{&I(X_1; Y_1) - I(X_{21} + X_{22}; Z_1), \\ &I(X_1; Y_1) - I(X_{21} + X_{22}; Z_2), \\ &I(X_{21} + X_{22}; Y_{21}, Y_{22}) - I(X_{21} + X_{22}; Z_1), \\ &I(X_{21} + X_{22}; Y_{21}, Y_{22}) - I(X_{21} + X_{22}; Z_2)\}. \end{aligned} \tag{35}$$

Based on the joint distribution of U and X, we obtain

$$\begin{aligned} I(X_1; Y_1) &\doteq \frac{1}{2}\log \text{SNR}, \\ I(X_{21} + X_{22}; Z_1) &\doteq I(X_{21} + X_{22}; Z_2) \doteq 0 \cdot \log \text{SNR}, \\ I(X_{21} + X_{22}; Y_{21}, Y_{22}) &\doteq \frac{1}{2}\log \text{SNR}, \\ I(X_{21} + X_{22}; Y_{21}, Y_{22}) &\doteq \frac{1}{2}\log \text{SNR}. \end{aligned} \tag{36}$$

Hence $R \doteq (1/2)\log \text{SNR}$, and Scheme 3 achieves

$$s.d.o.f. = 1. \tag{37}$$

Compared to Scheme 1 and Scheme 3 introduces extra randomness in the encoder by introducing a prefix channel $U \to X$, and hence achieves the optimal *s.d.o.f.* We also note that for Gaussian wiretap channels, including the single-input single-output channel studied in [15] and the multi-input multi-output channel studied in [16–18], the prefix channel is not necessary to achieve the secrecy capacity, that is, $U = X$. However, the prefix channel is necessary to achieve the optimal *s.d.o.f.* for the parallel Gaussian compound wiretap channel.

From Schemes 2 and 3, we also observe that introducing randomness either into the information source or into the encoder strictly improves the *s.d.o.f.* and hence improves the secrecy rate.

6. MIMO Compound Wiretap Channels

In this section, we consider the MIMO compound wiretap channel in which the transmitter, the receivers, and the eavesdroppers are equipped with multiple antennas. We let N_t denote the number of antennas of the transmitter, N_r denote the number of antennas of the receivers, and N_e denote the number of antennas of the eavesdroppers. We assume that all receivers have the same number of antennas and all eavesdroppers have the same number of antennas, but our analysis below is also applicable without this assumption.

The channel input-output relationship at one time instant is given by

$$\begin{aligned} \underline{Y}_j &= H_j \underline{X} + \underline{W}_j \quad \text{for } j = 1, \ldots, J, \\ \underline{Z}_k &= G_k \underline{X} + \underline{V}_k \quad \text{for } k = 1, \ldots, K, \end{aligned} \tag{38}$$

where H_j for $j = 1, \ldots, J$ and G_k for $k = 1, \ldots, K$ are fixed matrices, and $\underline{W}_1, \ldots, \underline{W}_J$ and $\underline{V}_1, \ldots, \underline{V}_K$ are i.i.d. Gaussian random vectors with identity covariance matrices. We assume that the channel input is subject to an average power constraint:

$$\frac{1}{n}\sum_{i=1}^{n} \mathrm{E}\left[\underline{X}_i^T \underline{X}_i\right] \le P, \tag{39}$$

where i is the symbol time (i.e., channel use) index.

In the following, we first study the degraded MIMO compound wiretap channel, and then study the general MIMO compound wiretap channel. We use the following notation associated with matrices. We use $A \succeq 0$ to indicate that A is a positive semidefinite matrix, $A \succ 0$ to indicate that A is a positive definite matrix, and $A \succeq B$ to indicate that $A - B$ is a positive semidefinite matrix. The symbols $\preceq$ and $\prec$ indicate the oppositive meanings to those of $\succeq$ and $\succ$, respectively.

6.1. Degraded MIMO Compound Wiretap Channels. As in [19], we define the MIMO compound wiretap channel to be *degraded* if for each (j, k) pair, there exists a matrix D_{jk} such that $D_{jk}H_j = G_k$ and $D_{jk}D_{jk}^T \preceq I$. It is easy to check that for each (j, k) pair, the channel satisfies the Markov chain relationship $\underline{X} \to \underline{Y}_j \to \underline{Z}_k$.

Theorem 5. *The secrecy capacity of the degraded MIMO compound wiretap channel is given by*

$$C = \max_{Q:Q\succeq 0,\ Tr(Q)\le P} \min_{j,k} \frac{1}{2}\log \frac{\left|I + H_j Q H_j^T\right|}{\left|I + G_k Q G_k^T\right|}. \tag{40}$$

Proof. We only need to show that the secrecy capacity is given by

$$\min_{j,k} \frac{1}{2}\log \frac{\left|I + H_j Q H_j^T\right|}{\left|I + G_k Q G_k^T\right|}, \tag{41}$$

if the input is subject to the covariance matrix constraint

$$\frac{1}{n}\sum_{i=1}^{n} K_{\underline{X}_i} \preceq Q, \tag{42}$$

where $K_{\underline{X}_i}$ denotes the covariance matrix of $\underline{X}_i$ at symbol time i. Theorem 5 then follows by maximizing (41) over all Q that satisfy the power constraint, that is, $Tr(Q) \le P$.

The achievability follows from Theorem 3 by choosing $X \sim \mathcal{N}(0, Q)$. The proof of the converse is relegated to Appendix C. □

6.2. General MIMO Compound Wiretap Channels. In this subsection, we study the general MIMO compound wiretap channel defined in (38), where we do not make the the degradedness assumption.

Based on Theorem 1 by choosing $U = X \sim \mathcal{N}(0, Q)$, it is easy to see that the following secrecy rate is achievable.

Lemma 1. *For the general MIMO compound wiretap channel, an achievable secrecy rate is given by*

$$R = \max_{Q:Q\geq 0, Tr(Q)\leq P} \min_{j,k} \frac{1}{2} \log \frac{\left| I + H_j Q H_j^T \right|}{\left| I + G_k Q G_k^T \right|}. \tag{43}$$

In general, the maximization problem in (43) is difficult to solve. To gain some insight, we study the *s.d.o.f.* defined as in (21), but with SNR $= P/N_t$.

We design the following beamforming scheme. Let $r = \text{Rank}(\sum_{j=1}^{J} H_j^T H_j)$ and $\{\underline{u}_1, \ldots, \underline{u}_r\}$ be the eigenvectors of $\sum_{j=1}^{J} H_j^T H_j$ that correspond to nonzero eigenvalues. These vectors are directions along which at least one receiver may receive input signals. In fact, if we let $\{\underline{u}_{j1}, \ldots, \underline{u}_{jr_j}\}$ be the eigenvectors of $H_j^T H_j$ that correspond to nonzero eigenvalues, then the vectors in the set $\{(\underline{u}_{j1}, \ldots, \underline{u}_{jr_j}) : j = 1, \ldots, J\}$ span the same vector space as $\{\underline{u}_1, \ldots, \underline{u}_r\}$.

We let $\{\underline{u}_{r+1}, \ldots, \underline{u}_{N_t}\}$ be the eigenvectors of $\sum_{j=1}^{J} H_j^T H_j$ that correspond to zero eigenvalues. We further let

$$U = [\underline{u}_1 \cdots \underline{u}_r], \qquad U^{\perp} = \left[\underline{u}_{r+1} \cdots \underline{u}_{N_t}\right]. \tag{44}$$

Then we have

$$\sum_{j=1}^{J} H_j^T H_j = [U \;\; U^{\perp}] \begin{bmatrix} \Lambda_r & \\ & 0_{N_t - r} \end{bmatrix} \begin{bmatrix} U^T \\ (U^{\perp})^T \end{bmatrix}, \tag{45}$$

where Λ_r denotes the diagonal matrix with the eigenvalues of $\sum_{j=1}^{J} H_j^T H_j$ as the diagonal components, and $0_{N_t - r}$ denotes the all-zero matrix of dimension $(N_t - r) \times (N_t - r)$.

We now let $\mathcal{L}$ be a subset of $\{1, 2, \ldots, r\}$ and assume $\mathcal{L} = \{l_1, \ldots, l_{|\mathcal{L}|}\}$, where $|\mathcal{L}|$ indicates the number of components in the set $\mathcal{L}$. We then let $\mathcal{L}^c$ denote the complement of $\mathcal{L}$ with respect to the set $\{1, 2, \ldots, r\}$ and assume $\mathcal{L}^c = \{l'_1, \ldots, l'_{r-|\mathcal{L}|}\}$. Let

$$U_{\mathcal{L}} = \left[\underline{u}_{l_1} \cdots \underline{u}_{l_{|\mathcal{L}|}}\right], \quad U_{\mathcal{L}^c} = \left[\underline{u}_{l'_1} \cdots \underline{u}_{l'_{r-|\mathcal{L}|}}\right]. \tag{46}$$

If we choose the beamforming directions to be column vectors in $U_{\mathcal{L}}$ and allocate power equally for these directions, then the input covariance matrix is given by

$$Q_{\mathcal{L}} = \frac{P}{|\mathcal{L}|} [U_{\mathcal{L}} \quad U_{\mathcal{L}^c} \quad U^{\perp}] \begin{bmatrix} I_{|\mathcal{L}|} & & \\ & 0 & \\ & & 0 \end{bmatrix} \begin{bmatrix} U_{\mathcal{L}}^T \\ U_{\mathcal{L}^c}^T \\ (U^{\perp})^T \end{bmatrix} \tag{47}$$

and we obtain

$$\begin{aligned} \left| I + H_j Q_{\mathcal{L}} H_j^T \right| &= \left| I + \frac{P}{|\mathcal{L}|} \left(H_j U_{\mathcal{L}}\right)^T \left(H_j U_{\mathcal{L}}\right) \right|, \\ \left| I + G_k Q_{\mathcal{L}} G_k^T \right| &= \left| I + \frac{P}{|\mathcal{L}|} (G_k U_{\mathcal{L}})^T (G_k U_{\mathcal{L}}) \right|. \end{aligned} \tag{48}$$

Hence we have

$$\begin{aligned} &\lim_{\text{SNR}\to\infty} \frac{(1/2) \log\left(\left| I + H_j Q H_j^T \right| / \left| I + G_k Q G_k^T \right| \right)}{(1/2) \log \text{SNR}} \\ &\quad = \text{Rank}\left(H_j U_{\mathcal{L}}\right) - \text{Rank}(G_k U_{\mathcal{L}}). \end{aligned} \tag{49}$$

Therefore, we have the following theorem.

Theorem 6. *An achievable secrecy degree of freedom of the MIMO compound wiretap channel is given by*

$$s.d.o.f. \geq \max_{\mathcal{L}} \min_{j,k} \left\{ \text{Rank}\left(H_j U_{\mathcal{L}}\right) - \text{Rank}(G_k U_{\mathcal{L}}) \right\}. \tag{50}$$

We note that each set $\mathcal{L}$ corresponds to one set of directions for which the transmitter allocates power, and hence corresponds to one power allocation strategy. The optimal achievable s.d.o.f. can be obtained by searching over all power allocation strategies. We note that $\text{Rank}(H_j U_{\mathcal{L}})$ *and* $\text{Rank}(G_k U_{\mathcal{L}})$ *in* (50) *can be interpreted as the dimensions of the projections of H_j and G_k, respectively, onto the vector space spanned by the column vectors of $U_{\mathcal{L}}$. Hence the achievable s.d.o.f. is determined by the geometry of the channel matrices to the receivers and eavesdroppers.*

For the special case $J = 1$, that is, there is only one receiver, the channel matrix to the receiver is H, and r becomes the rank of $H^T H$ and hence the rank of H. We should always choose $\mathcal{L} = \{1, \ldots, r\}$, and the resulting *s.d.o.f.* is given in the following corollary to Theorem 6.

Corollary 3. *For the MIMO compound wiretap channel with $J = 1$, an achievable secrecy degree of freedom is given by*

$$s.d.o.f. \geq \min_{k} \{Rank(H) - Rank(G_k U)\}, \tag{51}$$

where U is the matrix whose columns are the eigenvectors of $H^T H$ corresponding to nonzero eigenvalues.

We refer the reader to [7] for an example of MIMO compound wiretap channel for which particular signaling scheme transforms the channel into an equivalent parallel Gaussian compound wiretap channel, and a simple scheme can hence be constructed to achieve the *s.d.o.f.* for the channel.

7. Discussion and Conclusions

In this paper, we have studied the compound wiretap channel, which provides a general framework for examining multicast communication with multiple eavesdroppers. We have obtained lower and upper bounds on the secrecy capacity for the general compound wiretap channel and have established the secrecy capacity for the degraded and semideterministic channel. We have further obtained the secrecy capacity for the degraded parallel Gaussian and degraded MIMO compound wiretap channels. The secrecy rate/capacity in general has a worst-case interpretation.

We have also introduced the notion of the secrecy degree of freedom, which captures the most important factors that affect the scaling behavior of the secrecy capacity at high SNR. For the parallel Gaussian compound channel, we have demonstrated that the *s.d.o.f.* depends only on the maximal number of subchannels that one eavesdropper can access and does not depend on the number of eavesdroppers. We have also shown that randomizing either source information or the encoder strictly improves the *s.d.o.f.* for an example case when $J > 1$ and $K > 1$. For the MIMO compound wiretap channel, we have shown that the achievable *s.d.o.f.* is determined by the geometries of the matrices describing the channels to the receivers and eavesdroppers. We have also noted that it has been shown via a few example channels in [7] that there are simple schemes to achieve the *s.d.o.f.* in many cases.

We finally note that the capacity of the general compound wiretap channel is still not known. Several interesting special cases are worth addressing, including the Gaussian parallel compound wiretap channel with multiple receivers and multiple eavesdroppers and the general MIMO compound wiretap channel. Understanding the *s.d.o.f.* of these scenarios may be a useful first step. The techniques for studying the compound broadcast channel without secrecy constraints [20] may be useful here. In particular, designing a zero-forcing transmission scheme over multiple time slots for the MIMO compound wiretap channel as in [20] may be useful in studying the *s.d.o.f.* However, we remark that one cannot expect the eavesdropper to look only at subspaces. As a more general model, the MIMO compound broadcast channel is also interesting to study. Some recent studies [21] and [22] have provided useful techniques for further study of this topic.

Appendices

A. Proof of Theorem 1

The idea of the proof is to show there exists a codebook that consists of a number of subcodebooks (similar to [1]). Each receiver can successfully decode over the entire codebook, but all eavesdroppers can successfully decode only within each subcodebook. Hence the transmitter maps messages to different subcodebooks to confuse the eavesdroppers and achieve perfect secrecy.

For a given joint distribution $P_X P_{Y_1 \dots Y_J Z_1 \dots Z_K | X}$, it is sufficient to show the following rate is achievable:

$$R = \min_j I\left(X; Y_j\right) - \max_k I(X; Z_k). \tag{A.1}$$

Then the rate given in (6) is achievable by prefixing a discrete memoryless channel from U to X with the transition distribution $P_{X|U}$ to the transmitter (similar to [11, Lemma 4]).

We first prove a useful lemma that simplifies the proof later on.

Lemma A.1. *If $I(X; Z_1) < I(X; Z_2)$, then there exists a random variable $\widetilde{Z}$ such that $I(X; Z_1, \widetilde{Z}) = I(X; Z_2)$ and $\widetilde{Z}$ satisfies the Markov chain $X \to (Z_1, Z_2) \to \widetilde{Z}$.*

Proof. Let U be a binary random variable with distribution $Pr\{U = 1\} = p$ and $Pr\{U = 2\} = 1 - p$, and U is independent of all other random variables in the model under consideration. Let $\widetilde{Z} = (Z_U, U)$. Clearly, $\widetilde{Z}$ satisfies the Markov chain condition given in the lemma. Let

$$\begin{aligned} f(p) &= I\left(X; Z_1, \widetilde{Z}\right) = I(X; Z_1, Z_U, U) \\ &= I(X; Z_1, Z_U \mid U) \\ &= pI(X; Z_1) + (1 - p) I(X; Z_1, Z_2). \end{aligned} \tag{A.2}$$

It is clear that

$$\begin{aligned} f(1) &= I(X; Z_1) < I(X; Z_2), \\ f(0) &= I(X; Z_1 Z_2) \geq I(X; Z_2). \end{aligned} \tag{A.3}$$

Since $f(p)$ is a continuous function for $0 \leq p \leq 1$, there must exist p^* such that $f(p^*) = I(X; Z_2)$. Therefore, $\widetilde{Z} = (Z_U, U)$ with U having distribution $Pr\{U = 1\} = p^*$ satisfies $I(X; Z_1, \widetilde{Z}) = I(X; Z_2)$. □

Based on the previous lemma, it is sufficient to consider enhanced eavesdroppers, each with outputs $Z_k' = (Z_k, \widetilde{Z}_k)$ such that $I(X; Z_k') = \max_k I(X; Z_k)$. It is clear that if perfect secrecy can be achieved for the enhanced eavesdroppers, it must be achieved for the original eavesdroppers.

We now consider the following codebook:

$$\mathcal{C} = \left\{ x_{ab}^n,\ a = 1, \dots, 2^{nR};\ b = 1, \dots, 2^{n \max_k I(X; Z_k)} \right\}, \tag{A.4}$$

where R is given in (A.1). We assume all codewords are strongly typical, that is, $x_{ab}^n \in T_\epsilon^n(P_X)$, where $T_\epsilon^n(P_X)$ denotes the strongly jointly ϵ-typical set (see Section 1.2, [23]) based on the distribution P_X.

We define the following probabilities of error when the codeword x_{ab}^n is transmitted:

$$\begin{aligned} \lambda_{jab} &= \text{error probability for receiver } j \\ &\qquad \text{in determining } (a, b), \\ \eta_{kb|a} &= \text{error probability for eavesdropper } k \\ &\qquad \text{in determining } b \text{ given } a. \end{aligned} \tag{A.5}$$

Let p_{ab} be the probability with which codeword x_{ab}^n is transmitted. We further define the following average probabilities of error:

$$\begin{aligned} \lambda_j &= \sum_{ab} p_{ab} \lambda_{jab}, \\ \eta_k &= \sum_{ab} p_{ab} \lambda_{kb|a}. \end{aligned} \tag{A.6}$$

The following lemma guarantees existence of a certain codebook, which will be used for encoding.

Lemma A.2. *For any $0 < \epsilon < 1$, there exists a codebook as described in* (A.4)*, such that, for sufficiently large n,*

$$\begin{aligned} \lambda_j &< \epsilon \quad \text{for } j = 1, \dots, J, \\ \eta_k &< \epsilon \quad \text{for } k = 1, \dots, K. \end{aligned} \tag{A.7}$$

Proof. We prove the lemma by a random coding technique. We define the following sum of error probabilities:

$$p_e = \sum_j \lambda_j + \sum_k \eta_k = \sum_{jab} p_{ab}\lambda_{jab} + \sum_{kab} \eta_{kb|a}. \tag{A.8}$$

We show that the average of p_e over a random codebook ensemble is small for sufficiently large codeword length n. Then, there must exist at least one codebook such that p_e is small for sufficiently large n.

For a given distribution P_X, we generate codewords x_{ab}^n, each uniformly drawn from the set $T_\epsilon^n(P_X)$. Index x_{ab}^n via $a = 1,\ldots,2^{nR}$ and $b = 1,\ldots,2^{n\max_k I(X;Z_k)}$.

Suppose that the codeword x_{ab}^n is transmitted and define the following decoding strategies at the receivers and the eavesdroppers.

(1) Receiver j declares that the index pair of x_{ab}^n is $(\hat{a},\hat{b})$ if there is a unique index pair such that $(x_{\hat{a},\hat{b}}^n, y_j^n) \in T_\epsilon^n(P_{XY_j})$.

(2) Eavesdropper k, given the index a, declares that the index of x_{ab}^n is $\hat{b}$ if there is a unique index such that $(x_{a,\hat{b}}^n, z_k'^n) \in T_\epsilon^n(P_{XZ_k'})$, where Z_k' denotes the enhanced output.

We can compute $\mathrm{E}_{\mathbb{C}}[p_e]$ by following the standard techniques as in [24, Chapter 14], where $\mathrm{E}_{\mathbb{C}}$ indicates an average over the random codebook ensemble. We can show that

$$\mathrm{E}_{\mathbb{C}}[p_e] < \epsilon, \tag{A.9}$$

for sufficiently large codeword length n, based on the sizes of indices a and b.

Hence there exists one codebook such that for sufficiently large codebook size n

$$p_e = \sum_j \lambda_j + \sum_k \eta_k < \epsilon. \tag{A.10}$$

This leads to the conclusion that for sufficiently large codebook size n,

$$\lambda_j < \epsilon, \qquad \eta_k < \epsilon \tag{A.11}$$

for $j = 1,\ldots,J$ and $k = 1,\ldots,K$.

Based on the codebook that satisfies the property given in Lemma A.2, we define the encoding as follows. We map each message w to a codeword x_{wb}^n with b chosen uniformly over the set $\{1,\ldots,2^{n\max_k I(X;Z_k)}\}$. Based on Lemma A.2, it is clear that each receiver can decode the message W with small probability of error. For each enhanced eavesdropper, we follow steps similar to those in [10] to obtain the following equivocation rate:

$$\frac{1}{n} H\left(W \mid Z_k'^n\right) \geq R - \epsilon_1, \tag{A.12}$$

where $\epsilon_1 \to 0$ as $n \to \infty$. This concludes the proof. □

B. Proof of the Converse for Theorem 4

We follow steps that are similar to those given in [25] except for the step of single letter characterization. We include the proof here for the sake of completeness.

We consider a code with length n and average error probability P_e. The probability distribution we consider is

$$P_W P_{X^n|W} \prod_{i=1}^{n} \left[\prod_{k=1}^{K} P_{Y_i|X_i} P_{Z_{ki}|X_i}\right], \tag{B.1}$$

where $P_{Y_i|X_i}$ is a deterministic distribution, and thus takes values of only 0 or 1.

By Fano's inequality [24, Section 2.11], we have

$$H(W \mid Y^n) \leq nRP_e + 1 := n\delta, \tag{B.2}$$

where $\delta \to 0$ if $P_e \to 0$.

For each eavesdropper k, since we achieve perfect secrecy, we obtain the following bound:

$$\begin{aligned} nR = nR_e &\leq H\left(W \mid Z_k^n\right) \\ &= I\left(W; Y^n \mid Z_k^n\right) + H\left(W \mid Y^n, Z_k^n\right) \\ &\overset{(a)}{\leq} H\left(Y^n \mid Z_k^n\right) - H\left(Y^n \mid W, Z_k^n\right) + n\delta \\ &\leq H\left(Y^n \mid Z_k^n\right) + n\delta \\ &\overset{(b)}{\leq} \sum_{i=1}^{n} H(Y_i \mid Z_{k_i}) + n\delta, \end{aligned} \tag{B.3}$$

where (a) follows from Fano's inequality, and (b) follows from the chain rule and because conditioning does not increase entropy.

We now introduce a random variable Q that is independent of all other random variables in this model, and is uniformly distributed over $\{1,2,\ldots,n\}$. Define $X = X_Q$, $Y = Y_Q$, and $Z_k = Z_{kQ}$. It is clear that these random variables satisfy the Markov chain condition $Q \to X \to (Y, Z_k)$. Using these definitions, (B.3) becomes

$$\begin{aligned} R &\leq H(Y_Q \mid Z_{kQ}, Q) + \delta \\ &\leq H(Y_Q \mid Z_{kQ}) + \delta \\ &= H(Y \mid Z_k) + \delta. \end{aligned} \tag{B.4}$$

The bound given in (B.4) is applicable for $k = 1,\ldots,K$, and hence we obtain

$$R \leq \min_k H(Y \mid Z_k) + \delta \tag{B.5}$$

which completes the proof.

C. Proof of the Converse for Theorem 5

We first prove the following lemma, which gives two useful properties.

Lemma C.1. *Consider matrices D, H, G, and X of dimensions conformal with the following operations. If $DH = G$ and $DD^T \preceq I$, then $f(X) = \log(|I + HXH^T|/|I + GXG^T|)$ is a concave function of $X \succeq 0$. Furthermore, $f(X) \leq f(X + \Delta)$ if $X \succeq 0$, and $\Delta \succeq 0$.*

Proof.

$$f(X) = \log \frac{\left|I + HXH^T\right|}{|I + HXH^T D^T D|} = \log \frac{\left|I + HXH^T\right|}{\left|\left((D^T D)^{-1} - I\right) + I + HXH^T\right| |D^T D|}. \tag{C.1}$$

By [26, Lemma II.3], the preceding function is concave in $I + HXH^T$ and is hence concave in X.

To show the second property, we recall the following property given in [27, page 3942]. If $A \succeq 0, \Delta \succeq 0$ and $B \succ 0$, then

$$\frac{|B|}{|A + B|} \leq \frac{|B + \Delta|}{|A + B + \Delta|}. \tag{C.2}$$

Applying the above property to (C.1), we obtain

$$f(X) \leq \log \frac{\left|I + H(X + \Delta)H^T\right|}{\left|\left((D^T D)^{-1} - I\right) + I + H(X + \Delta)H^T\right| |D^T D|} = \log \frac{\left|I + H(X + \Delta)H^T\right|}{|I + G(X + \Delta)G^T|} = f(X + \Delta). \tag{C.3}$$

□

To prove the converse, we first have the following bound for any (j, k) pair by referring to [15, Section III]:

$$R = R_e \leq \frac{1}{n}\sum_{i=1}^{n} I\left(\underline{X}_i; Y_{j,i} \mid \underline{Z}_{k,i}\right) \leq \frac{1}{n}\sum_{i=1}^{n}\left[h\left(\underline{Y}_{j,i} \mid \underline{Z}_{k,i}\right) - h\left(\underline{Y}_{j,i} \mid \underline{X}_i, \underline{Z}_{k,i}\right)\right] = \frac{1}{n}\sum_{i=1}^{n}\left[h\left(\underline{Y}_{j,i} \mid \underline{Z}_{k,i}\right) - h\left(\underline{W}_{j,i} \mid \mathrm{V}_{k,i}\right)\right]. \tag{C.4}$$

It is easy to see that the second term is independent of the distribution of $\underline{X}_i$. The first term is maximized by Gaussian $\underline{X}_i$ if the covariance matrix of $\underline{X}_i$ is fixed to be $K_{\underline{X}_i}$. This is because $h(\underline{Y}_{j,i} \mid \underline{Z}_{k,i})$ is maximized by jointly Gaussian $\underline{Y}_{j,i}$ and $\underline{Z}_{k,i}$ for a fixed covariance matrix $Q_{\underline{Y}_{j,i}\underline{Z}_{k,i}}$. Therefore, we have the following bound:

$$R_e \leq \frac{1}{n}\sum_{i=1}^{n}\frac{1}{2}\log\frac{\left|I + HK_{\underline{X}_i}H^T\right|}{\left|I + GK_{\underline{X}_i}G^T\right|} \overset{(a)}{\leq} \frac{1}{2}\log\frac{\left|I + H\left((1/n)\sum_{i=1}^{n} K_{\underline{X}_i}\right)H^T\right|}{\left|I + G\left((1/n)\sum_{i=1}^{n} K_{\underline{X}_i}\right)G^T\right|} \overset{(b)}{\leq} \frac{1}{2}\log\frac{\left|I + HQH^T\right|}{|I + GQG^T|}, \tag{C.5}$$

where (a) follows from the degradedness assumption and the concavity property given in Lemma C.1, and (b) follows because $(1/n)\sum_{i=1}^{n} K_{\underline{X}_i} \preceq Q$ and from the monotonicity property given in Lemma C.1.

Acknowledgments

The material in this paper has been presented in part at the 45th Annual Allerton Conference on Communication, Control, and Computing, Monticello, IL, USA, Sept. 2007, and in part at the Annual IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC), Cannes, France, Sept. 2008. The work of Y. Liang was supported by the National Science Foundation CAREER Award under Grant CCF-08-46028. The work of G. Kramer was supported in part by the Board of Trustees of the University of Illinois Subaward no. 04-217 under NSF Grant CCR-0325673 and the Army Research Office under ARO Grant W911NF-06-1-0182. The work of H. V. Poor was supported by the National Science Foundation under Grants ANI-03-38807, CNS-06-25637 and CCF-07-28208. The work of S. Shamai was supported by the European Commission in the framework of the FP7 Network of Excellence in Wireless Communications NEWCOM++.

References

[1] A. D. Wyner, "The wire-tap channel," *Bell System Technical Journal*, vol. 54, no. 8, pp. 1355–1387, 1975.

[2] H. Yamamoto, "Coding theorem for secret sharing communication systems with two noisy channels," *IEEE Transactions on Information Theory*, vol. 35, no. 3, pp. 572–578, 1989.

[3] H. Yamamoto, "A coding theorem for secret sharing communication systems with two Gaussian wiretap channels," *IEEE Transactions on Information Theory*, vol. 37, no. 3, part 1, pp. 634–638, 1991.

[4] P. Wang, G. Yu, and Z. Zhang, "On the secrecy capacity of fading wireless channel with multiple eavesdroppers," in *Proceedings of the IEEE International Symposium on Information Theory (ISIT '07)*, pp. 1301–1305, Nice, France, June 2007.

[5] A. Khisti, A. Tchamkerten, and G. Wornell, "Secure broadcasting over fading channels," *IEEE Transactions on Information Theory*, vol. 54, no. 6, pp. 2453–2469, 2008.

[6] N. Cai and R. W. Yeung, "Secure network coding," in *Proceedings of the IEEE International Symposium on Information Theory (ISIT '02)*, p. 323, Lausanne, Switzerland, June-July 2002.

[7] Y. Liang, G. Kramer, H. V. Poor, and S. Shamai (Shitz), "Compound wire-tap channels," in *Proceedings of the 45th Annual Allerton Conference on Communication, Control, and Computing*, Monticello, IL, USA, September 2007.

[8] T. Liu, V. Prabhakaran, and S. Vishwanath, "The secrecy capacity of a class of parallel Gaussian compound wiretap channels," in *Proceedings of IEEE International Symposium on Information Theory (ISIT '08)*, pp. 116–120, Toronto, Canada, July 2008.

[9] Y. Liang, H. V. Poor, and S. Shamai (Shitz), "Information theoretic security," *Foundations and Trends in Communications and Information Theory*, vol. 5, no. 4-5, pp. 355–580, 2008.

[10] Y. Liang and H. V. Poor, "Multiple-access channels with confidential messages," *IEEE Transactions on Information Theory*, vol. 54, no. 3, pp. 976–1002, 2008.

[11] I. Csiszár and J. Körner, "Broadcast channels with confidential messages," *IEEE Transactions on Information Theory*, vol. 24, no. 3, pp. 339–348, 1978.

[12] Y. Liang, H. V. Poor, and S. Shamai (Shitz), "Secure communication over fading channels," *IEEE Transactions on Information Theory, Special Issue on Information Theoretic Security*, vol. 54, no. 6, pp. 2470–2492, 2008.

[13] C. Mitrpant, A. J. H. Vinck, and Y. Luo, "An achievable region for the Gaussian wiretap channel with side information," *IEEE Transactions on Information Theory*, vol. 52, no. 5, pp. 2181–2190, 2006.

[14] D. Tse, "A deterministic model for wireless channels and its applications," in *Proceedings of the IEEE Information Theory Workshop (ITW '07)*, p. 607, Lake Tahoe, Calif, USA, September 2007.

[15] S. K. Leung-Yan-Cheong and M. E. Hellman, "The Gaussian wire-tap channel," *IEEE Transactions on Information Theory*, vol. 24, no. 4, pp. 451–456, 1978.

[16] A. Khisti and G. Wornell, "The MIMOME channel," in *Proceedings of the 45th Annual Allerton Conference on Communication, Control, and Computing*, Monticello, IL, USA, September 2007.

[17] F. Oggier and B. Hassibi, "The secrecy capacity of the MIMO wire-tap channel," in *Proceedings of the 45th Annual Allerton Conference on Communication, Control, and Computing*, Monticello, IL, USA, September 2007.

[18] T. Liu and S. Shamai (Shitz), "A note on the secrecy capacity of the multiple-antenna wiretap channel," *IEEE Transactions on Information Theory*, vol. 55, no. 6, pp. 2547–2553, 2009.

[19] H. Weingarten, T. Liu, S. Shamai (Shitz), Y. Steinberg, and P. Viswanath, "The capacity region of the degraded MIMO compound broadcast channel," in *Proceedings of the IEEE International Symposium on Information Theory (ISIT '07)*, pp. 766–770, Nice, France, June 2007.

[20] H. Weingarten, S. Shamai (Shitz), and G. Kramer, "On the compound MIMO broadcast channel," in *Proceedings of the Information Theory and Applications Workshop (ITA '07)*, La Jolla, Calif, USA, January-February 2007.

[21] M. Kobayashi, Y. Liang, S. Shamai (Shitz), and M. Debbah, "On the compound MIMO broadcast channels with confidential messages," in *Proceedings of the IEEE International Symposium on Information Theory (ISIT '09)*, pp. 1283–1287, Seoul, Korea, June-July 2009.

[22] G. Bagheri-Karam, A. Motahari, and A. K. Khandani, "The secrecy capacity region of the degraded vector Gaussian broadcast channel," in *Proceedings of the IEEE International Symposium on Information Theory (ISIT '09)*, pp. 2772–2776, Seoul, Korea, June-July 2009.

[23] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*, Akadémiai Kiadó, Budapest, Hungary, 1981.

[24] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, John Wiley & Sons, New York, NY, USA, 1991.

[25] J. Grubb, S. Vishwanath, Y. Liang, and H. V. Poor, "Secrecy capacity of semi-deterministic wire-tap channels," in *Proceedings of the IEEE Information Theory Workshop on Information Theory for Wireless Networks (ITW '07)*, pp. 199–202, July 2007.

[26] S. N. Diggavi and T. M. Cover, "The worst additive noise under a covariance constraint," *IEEE Transactions on Information Theory*, vol. 47, no. 7, pp. 3072–3081, 2001.

[27] H. Weingarten, Y. Steinberg, and S. Shamai (Shitz), "The capacity region of the Gaussian multiple-input multiple-output broadcast channel," *IEEE Transactions on Information Theory*, vol. 52, no. 9, pp. 3936–3964, 2006.